

# Notions from real analysis – a brief review

Stilian A. Stoev  
*Lecture Notes for STAT 520*

Fall 2009

## Contents

<b>1 Preliminaries: sets, functions and Cartesian products</b>	
<b>2 The Real line <math>\mathbb{R}</math>. Sequences, convergence in <math>\mathbb{R}</math> and <math>\mathbb{R}^k</math></b>	<b>5</b>
<b>3 Metric spaces: convergence and continuity. Topology</b>	<b>7</b>
<b>4 Functions on <math>\mathbb{R}^k</math>: continuity and differentiability</b>	<b>12</b>
<b>5 Riemann integration</b>	<b>15</b>
<b>6 Miscellanea</b>	<b>18</b>
<b>7 Exercises</b>	<b>19</b>

## 1 Preliminaries: sets, functions and Cartesian products

A collection of arbitrary objects is called a *set*. The members of a set (if any) are called its elements or points.

- *Relations and operations with sets*

(belongs) For a given set  $A$ , if the object  $a$  is an element of  $A$ , we write “ $a \in A$ ” and say  $a$  belongs to  $A$ ; otherwise, we write  $a \notin A$  and say  $a$  does not belong to  $A$ .

(subset) A given set  $A$  is a sub-set of a set  $B$  if all elements of  $A$  are elements of  $B$ , denoted  $A \subset B$ .

(equality)  $A = B$  if  $A \subset B$  and  $B \subset A$ .

(intersection)  $A \cap B$  – the set comprised with the common elements of  $A$  and  $B$ .

(union)  $A \cup B$  – the set comprised with points that are elements of  $A$  or  $B$ .

(difference)  $A \setminus B$  – the set of all elements in  $A$  that are not elements of  $B$ .

(symmetric difference)  $A\Delta B := (A \setminus B) \cup (B \setminus A)$ .

We have:

$$(A \cup B) \cup C = A \cup (B \cup C) \quad \text{and} \quad (A \cap B) \cap C = A \cap (B \cap C),$$
$$(A \cup B) \cap C = (A \cap C) \cup (B \cap C) \quad \text{and} \quad (A \cap B) \cup C = (A \cup C) \cap (B \cup C).$$

In general, if  $A_\tau$ ,  $\alpha \in \mathcal{T}$ , is a collection of sets indexed by another set  $\mathcal{T}$ , we have:

$$\bigcup_{\tau \in \mathcal{T}} A_\tau := \{x : x \in A_\tau \text{ for some } \tau \in \mathcal{T}\} \quad \text{and} \quad \bigcap_{\tau \in \mathcal{T}} A_\tau := \{x : x \in A_\tau \text{ for all } \tau \in \mathcal{T}\}.$$

The index set can be finite:  $\{1, 2, \dots, n\}$ ; countably infinite:  $\mathcal{T} = \{1, 2, \dots\}$  or uncountably infinite.

• *Some more notation:*

$\emptyset$  – the special set with no elements, called the *empty set*.

$\mathbb{N}$  – the set of *natural numbers*:  $\{1, 2, \dots\}$

$\mathbb{Z}$  – the set of *integer numbers*:  $\{\dots, -1, 0, 1, 2, \dots\}$

$\mathbb{Q}$  – the set of *rational numbers*:  $\{p/q : p, q \in \mathbb{Z}, q \neq 0\}$ .

$\mathbb{R}$  – the set of *real numbers*:  $(-\infty, \infty)$ .

$\mathbb{C}$  – the set of *complex numbers*.

$2^A$  – for a given set  $A$ , the set  $2^A$  is the set of all sub-sets of  $A$ , called the *power set* of  $A$ .

We often specify a set as follows. If  $a < b$  are two real numbers, then

$$A = \{x : a < x < b, x \text{ – real number}\},$$

denotes the set of real numbers between  $a$  and  $b$ , i.e. the interval  $(a, b)$ .

See the exercises in Section 7 for some useful facts on working with sets.

• *Russell's paradox*

One cannot define sets in an arbitrary way by using common logic. Indeed consider the set:

$$A = \{\text{all sets } B \text{ that are not members of themselves}\}.$$

*Is the set  $A$  a member of itself?* If “yes”, then  $A$  cannot be among the “sets that are not members of themselves”, but this contradicts the definition of  $A$  since  $A$  is a member of itself. If “no”, then  $A$  is a member of itself, according to the way  $A$  is defined, which also leads to a contradiction.

Such “recursive” definitions of sets should be avoided. This paradox is resolved by restricting the definition of sets to all subsets of a *very large* and *fixed* set  $X$ . In our considerations, we will not encounter the Russell's paradox.

• *Functions between sets*

Let  $A$  and  $B$  be two non-empty sets. A correspondence  $f$  assigning to any element  $a$  of  $A$  a *unique* element  $b = f(a)$  of  $B$  is said to be a function, notation

$$f : A \rightarrow B \quad \text{or} \quad f : \begin{cases} A \rightarrow B \\ a \mapsto f(a) \end{cases}$$

Domain of  $f$ , denoted  $\text{dom}(f)$  is the largest sub-set where  $f$  is defined. So if  $f : A \rightarrow B$ , then  $\text{dom}(f) = A$ .

Range of  $f$ , denoted  $\text{range}(f) = \{f(a) : \forall a \in \text{dom}(f)\}$  is the set of all possible values of  $f$ . Often, if  $f : A \rightarrow B$ , then  $\text{range}(f)$  is denoted  $f(A)$ .

Onto: A map  $f : A \rightarrow B$  is *onto* or *surjective* if  $f(A) = B$ .

1-to-1: A map  $f : A \rightarrow B$  is *1-to-1* or *injective* if  $f(a_1) = f(a_2)$  implies  $a_1 = a_2$ , i.e. different  $a$ 's are mapped into different  $f(a)$ 's.

Bijection: A map  $f : A \rightarrow B$  is a *bijection* if it is 1-to-1 and onto.

Pre-image: For a map  $f : A \rightarrow B$  and a set  $C \subset B$ , the pre-image of  $C$  is  $f^{-1}(C) := \{a \in A : f(a) \in C\}$ .

Notice that if  $f : A \rightarrow B$  is 1-to-1, then one can define the *inverse map*  $f^{-1} : \text{range}(f) \rightarrow A$ , so that  $f^{-1}(f(a)) = a$ . Note that the domain of  $f^{-1}$  is the range of  $A$ . If in addition  $f$  is onto then  $B = f(A)$  and  $f^{-1} : B \rightarrow A$  is also a bijection. Observe that the inverse map  $f^{-1}$ , whenever defined, is always 1-to-1!

**Examples:**

1.  $f : \mathbb{R} \rightarrow \mathbb{R}, x \mapsto f(x) = x^2$
2.  $f : \mathbb{N} \rightarrow \{0, 1\}, f(x) = 0$ , if  $x$  - even  $f(x) = 1$ , if  $x$  - odd.
3.  $f : A \rightarrow A, x \mapsto f(x) = x$  is the special map called the *identity* and denoted  $\text{id}_A$ .
4. Let  $X \neq \emptyset$ , and  $A \subset X$ . The *indicator function* of the set  $A$ , denoted  $I_A : X \rightarrow \{0, 1\}$ , is such that

$$I_A(x) := \begin{cases} 1 & , \text{ if } x \in A \\ 0 & , \text{ otherwise} \end{cases}$$

5.  $f : A \rightarrow B$  and  $g : B \rightarrow C$ , then  $g \circ f : A \rightarrow C$ , the composition of  $f$  and  $g$ , is defined as:  $x \mapsto g(f(x)), \forall x \in A$ . Notice that the order of  $f$  and  $g$  is important!

• *Properties of the unions and intersections when applying maps and taking pre-images*

**Proposition 1.1** Let  $f : X \rightarrow Y$ , be a an arbitrary map. Consider collections of sets  $A_\tau \subset X, \tau \in \mathcal{T}$  and  $B_\tau \subset f(Y), \tau \in \mathcal{T}$ . Then:

- a.  $f(\cup_\tau A_\tau) = \cup_\tau f(A_\tau)$
- b.  $f(\cap_\tau A_\tau) \subset \cap_\tau f(A_\tau)$
- c.  $f^{-1}(\cap_\tau B_\tau) = \cap_\tau f^{-1}(B_\tau)$  and  $f^{-1}(\cup_\tau B_\tau) = \cup_\tau f^{-1}(B_\tau)$ .
- d.  $f(C) \setminus f(A) \subset f(C \setminus A)$  and  $f^{-1}(D) \setminus f^{-1}(E) \equiv f^{-1}(D \setminus E)$ .
- e.  $f(f^{-1}(A)) \subset A$  and  $f^{-1}(f(A)) \supset A$ .
- f. If  $f$  is 1-to-1, then  $f(\cap_\tau A_\tau) = \cap_\tau f(A_\tau)$ .

• *Cartesian products. Powers of sets.*

Let  $A$  and  $B$  be two non-empty sets. The Cartesian product  $A \times B$  of the sets  $A$  and  $B$  is defined as:

$$A \times B := \{(a, b) : a \in A, b \in B\}.$$

That is, the set of all *ordered pairs*  $(a, b)$ , where the first is element from  $A$  and the second from  $B$ .

One easily extends the definition of a Cartesian product to any finite number of terms:

$$A_1 \times \cdots \times A_k := \{(a_1, \dots, a_k) : a_i \in A_i, i = 1, \dots, k\}, \quad k \in \mathbb{N},$$

where  $A_1, \dots, A_k$  are non-empty. One can easily see that there is a natural correspondence between  $(A \times B) \times C$  and  $A \times (B \times C)$ , so that one can identify the two.

We next extend the notion of a Cartesian product to the case of infinite number of terms. Let  $A_\tau$ ,  $\tau \in \mathcal{T}$  be a collection of non-empty sets, indexed by a non-empty index set  $\mathcal{T}$  (which can be finite or infinite). Define

$$\times_{\tau \in \mathcal{T}} A_\tau := \{f : \mathcal{T} \rightarrow \cup_{\tau \in \mathcal{T}} A_\tau : f(\tau) \in A_\tau\}.$$

If  $A_\tau = A$ ,  $\forall \tau \in \mathcal{T}$  and  $\mathcal{T} = \mathbb{N}$ , then

$$\times_{\tau \in \mathbb{N}} A_\tau = A^{\mathbb{N}} = \{(a_1, a_2, \dots) \mid a_i \in A, i \in \mathbb{N}\}$$

is the set of all infinite sequences with elements of  $A$ . Indeed, any function  $f : \mathbb{N} \rightarrow A$  corresponds to a sequence  $a_i := f(i)$ ,  $i \in \mathbb{N}$  and conversely.

**Examples:**

1. Let  $A = \{1, 2, \dots, n\}$  and  $B = \{x_1, \dots, x_m\}$ . Then  $A \times B$  consists of  $nm$  elements of the type:  $(i, x_j)$ ,  $1 \leq i \leq n$ ,  $1 \leq j \leq m$ .
2. Consider the set  $X$  of all functions  $f : \mathbb{R} \rightarrow \mathbb{R}$ . According to the definition of a Cartesian product, we have that

$$X = \{f : \mathbb{R} \rightarrow \mathbb{R}\} \equiv \mathbb{R}^{\mathbb{R}} = \times_{\tau \in \mathbb{R}} \mathbb{R}.$$

Any function  $f$  corresponds to a “selection” of a unique element from  $\mathbb{R}$  for any index value  $\tau \in \mathbb{R}$  and hence can be treated as an element of the set  $\mathbb{R}^{\mathbb{R}}$ .

3. As indicated above, all functions  $f : \mathbb{N} \rightarrow \mathbb{R}$  are the elements of the Cartesian product  $\mathbb{R}^{\mathbb{N}} = \times_{\tau \in \mathbb{N}} \mathbb{R}$ .

Consider the Cartesian product  $\mathbb{R}^{\mathbb{N}}$ . Natural *projection maps* relate the Cartesian product spaces  $\mathbb{R}^{\mathbb{N}}$  and  $\mathbb{R}^n \equiv \mathbb{R} \times \cdots \times \mathbb{R}$ . Indeed, fix the dimension indices  $k_1 < k_2 < \cdots < k_n$ , and define:

$$\pi := \pi_{k_1, \dots, k_n} : \mathbb{R}^{\mathbb{N}} \rightarrow \mathbb{R}^n$$

as follows. For any  $f \in \mathbb{R}^{\mathbb{N}}$ , let  $g := \pi(f) : \{1, \dots, n\} \rightarrow \mathbb{R}$ , be such that  $g(i) = f(k_i)$ . That is,  $g$  picks out only the values of  $f$  at the dimensions  $k_1, \dots, k_n$ .

• *Sequences of sets. Limsup, liminf and limits of sets.*

Let  $A_n$ ,  $n \in \mathbb{N}$  be an arbitrary sequence of sets. We define the limsup and liminf of  $\{A_n\}$  as follows:

$$\limsup\{A_n\} := \bigcap_{m \geq 1} \left( \bigcup_{n \geq m} A_n \right) \quad \text{and} \quad \liminf\{A_n\} := \bigcup_{m \geq 1} \left( \bigcap_{n \geq m} A_n \right).$$

Interpretation:  $\limsup\{A_n\}$  coincides with the set

$$A^* := \{x : x \text{ belongs to infinitely many } A_n\text{'s}\}$$

on the other hand,  $\liminf\{A_n\}$ , is precisely the set

$$A_* := \{x : x \text{ belongs to all } A_n\text{'s for all sufficiently large } n\text{'s}\}.$$

One often uses the notation  $\overline{\lim}$  and  $\underline{\lim}$  for limsup and liminf, respectively.

If  $\limsup(A_n) = \liminf(A_n)$ , we then say that the sequence of sets  $A_n$ ,  $n \in \mathbb{N}$  has a limit, denoted:

$$\lim_n(A_n) := \limsup(A_n) = \liminf(A_n).$$

## 2 The Real line $\mathbb{R}$ . Sequences, convergence in $\mathbb{R}$ and $\mathbb{R}^k$

We take for granted the definition of the set of real numbers  $\mathbb{R} = (-\infty, \infty)$  and the usual operations and relations. We shall frequently use the notation:

$(a, b)$  – the open interval  $\{x \in \mathbb{R} : a < x < b\}$

$(a, b]$  – the semi-open interval  $\{x \in \mathbb{R} : a < x \leq b\}$

$[a, b)$  – the semi-open interval  $\{x \in \mathbb{R} : a \leq x < b\}$

$[a, b]$  – the closed interval  $\{x \in \mathbb{R} : a \leq x \leq b\}$

A non-empty set of real numbers  $A \subset \mathbb{R}$  is said to be *bounded below* if:

$$\exists x \in \mathbb{R} : x \leq a, \forall a \in A,$$

i.e. if  $x \leq A$ , for some  $x \in \mathbb{R}$  called a *lower bound* of  $A$ . If for some  $y \in \mathbb{R}$ ,  $a \leq y, \forall a \in A$ , then  $A$  is said to be *bounded above*, in this case,  $y$  is called an *upper bound* of  $A$ . If  $A \subset \mathbb{R}$  is bounded above and below, then it is said to be *bounded*.

Consider a non-empty set  $A \subset \mathbb{R}$  which is bounded below. The infimum of  $A$  is defined as the *greatest lower bound* of  $A$ , and denoted  $\inf A$ . That is, for any  $x \leq A$ , we have  $x \leq \inf A$ . We similarly define  $\sup A$  as the *least upper bound* of a bounded above set  $A$ .

**The completeness axiom:** Any bounded below (above) non-empty set  $A \subset \mathbb{R}$  has infimum (supremum, respectively).

One can show that if  $y = \inf A$  exists, then:

$$\forall \epsilon > 0, \exists a_\epsilon \in A, \text{ such that } a_\epsilon < y + \epsilon.$$

A similar (dual) property is valid for the supremum.

The completeness axiom implies that the Real line has no “holes”. That is, if  $L \cup U = \mathbb{R}$ ,  $L \cap U = \emptyset$ , and  $x < y$ , for any  $x \in L$  and  $y \in U$ , then either  $\sup L \in L$  or  $\inf U \in U$ .

**Examples:**

1. The set  $A = \{1/n : n = 1, 2, \dots\}$  is bounded with  $\sup A = 1$  and  $\inf A = 0$ .
2. The set  $A = \{\tan(n) : n \in \mathbb{N}\}$  does not have neither a lower nor an upper bound. Why?
3. Let  $A = \{x \in \mathbb{Q} : x > \pi\}$ . Then  $\inf A = \pi$ . Indeed,  $\pi \leq \inf A$ , but also for any  $\epsilon > 0$ , there is a rational number  $x = p/q$ ,  $p, q \in \mathbb{N}$  in the interval  $(\pi, \pi + \epsilon)$ .

**Definition 2.1** A sequence  $\{x_n\}_{n \in \mathbb{N}}$  of real numbers is said to be convergent with limit  $x \in \mathbb{R}$ , if:

$$\forall \epsilon, \exists N = N_\epsilon, \text{ such that } |x_n - x| < \epsilon, \forall n \geq N_\epsilon.$$

We write  $x_n \rightarrow x$ ,  $n \rightarrow \infty$  or  $\lim_{n \rightarrow \infty} x_n = x$ .

A point  $y \in \mathbb{R}$  is said to be a *limit point* of a sequence  $\{x_n\}_{n \in \mathbb{N}}$ , if  $x_{n_k} \rightarrow y$ ,  $k \rightarrow \infty$ , for some sub-sequence  $n_k \rightarrow \infty$  of integers.

**Theorem 2.1 (Bolzano–Weierstrass)** *If  $\{x_n\}_{n \in \mathbb{N}}$  is a bounded sequence (below and above), then it has at least one limit point.*

One can show that a sequence is convergent if and only if it has a unique limit point. Observe that monotone sequences have at most one limit point. Hence any monotone increasing (decreasing) sequences is convergent, provided that it is bounded.

We next define the liminf and limsup of a sequence, which yield the smallest and largest limit points of a sequence (when the latter are finite).

**Definition 2.2** Let  $\{x_n\}_{n \in \mathbb{N}}$  be bounded below. Then,

$$\liminf_{n \rightarrow \infty} x_n := \lim_{n \rightarrow \infty} \inf_{m \geq n} x_m$$

is the liminf of the sequence  $\{x_n\}$ . Similarly,  $\limsup_{n \rightarrow \infty} x_n := \lim_{n \rightarrow \infty} \sup_{m \geq n} x_m$ , is the limsup of the sequence  $\{x_n\}$ , when it is bounded above.

Clearly  $\liminf_{n \rightarrow \infty} x_n \leq \limsup_{n \rightarrow \infty} x_n$ . One can show that  $x_n \rightarrow x$ ,  $n \rightarrow \infty$ , if and only if

$$\liminf_{n \rightarrow \infty} x_n = \limsup_{n \rightarrow \infty} x_n. \tag{2.1}$$

In this case the last liminf and limsup are equal to  $x = \lim_{n \rightarrow \infty} x_n$ .

An important notion is that of a *Cauchy sequence*.

**Definition 2.3** The real sequence  $\{x_n\}_{n \in \mathbb{N}}$  is said to be a *Cauchy sequence* if

$$\forall \epsilon > 0, \exists N = N_\epsilon, \text{ such that } |x_n - x_m| < \epsilon, \forall m, n \geq N_\epsilon,$$

that is,  $|x_n - x_m| \rightarrow 0$ , as  $n, m \rightarrow \infty$ .

Any convergent sequence is a Cauchy sequence. One can also show that any Cauchy sequence of real numbers is convergent. We will see below that this is not always the case in general metric spaces.

• *Sequences and convergence in  $\mathbb{R}^k$*

An infinite sequence  $x_n = (x_n(i))_{i=1}^k \in \mathbb{R}^k$ ,  $n = 1, 2, 3, \dots$  converges to  $x = (x(i))_{i=1}^k \in \mathbb{R}^k$  if it converges component-wise:

$$\lim_{n \rightarrow \infty} x_n(i) = x(i), \quad \forall i = 1, \dots, k.$$

We then also write  $\lim_{n \rightarrow \infty} x_n = x$  or  $x_n \rightarrow x$ ,  $n \rightarrow \infty$ . One defines as in  $\mathbb{R}$  limit points of a sequence in  $\mathbb{R}^k$ . A set  $A \subset \mathbb{R}^k$  is *bounded* if it is bounded in each one of its  $k$  coordinates, i.e. its projections  $\pi_i(A)$  are bounded in  $\mathbb{R}$ , for all  $1 \leq i \leq k$ .

**Exercise:** Show that the Bolzano–Weierstrass Theorem remains valid in  $\mathbb{R}^k$ .

### 3 Metric spaces: convergence and continuity. Topology

**Definition 3.1** Let  $X$  be a non-empty set. A function  $\rho : X \times X \rightarrow [0, \infty)$  is said to be a metric if the following three conditions hold:

- i.  $\rho(x, y) = 0$ , if and only if  $x = y$ .
- ii.  $\rho(x, y) = \rho(y, x)$ ,  $\forall x, y \in X$
- iii.  $\rho(x, y) \leq \rho(x, z) + \rho(z, y)$ ,  $\forall x, y, z \in X$ .

The pair  $(X, \rho)$  of the set  $X$  equipped with the metric  $\rho$  is said to be a *metric space*.

**Examples:**

1. The real line  $\mathbb{R}$  is a metric space with the metric  $\rho(x, y) := |x - y|$ .
2. The Euclidean space  $\mathbb{R}^n$ ,  $n \geq 2$  equipped with the metric  $\rho(x, y) := \|x - y\|_2$  is a metric space.
3. Let  $Y := \{1, 2, \dots, n\}$  and let  $X := 2^Y$  be the power set of  $Y$ , that is, the set of all sub-sets of  $Y$ . For any two subsets  $A, B \subset Y$ , let

$$\rho(A, B) := |A \Delta B|, \tag{3.1}$$

where  $|A|$  denotes the number of elements of the set  $A$ .

Then,  $(X, \rho)$  is a metric space. Indeed, conditions **i.** and **ii.** are straightforward. The triangle inequality follows from the fact that

$$\rho(A, B) = \sum_{x \in Y} I_{A \Delta B}(x) = \sum_{x \in Y} |I_A(x) - I_B(x)|,$$

and then by using the triangle inequality.

This is the so-called Hamming distance between the characteristic strings of 0's and 1's representing the subsets of  $Y$ .

4. Let  $X = \ell_2 = \{(a_j)_{j=1}^\infty : \sum_{j=1}^\infty a_j^2 < \infty, a_j \in \mathbb{R}\}$  and let  $\rho(a, b) := (\sum_{j=1}^\infty (a_j - b_j)^2)^{1/2}$ . Then, one can show that  $(X, \rho)$  is a metric space.

5. Let  $X = \{f : [0, 1] \rightarrow \mathbb{R} \mid f \text{ is continuous}\}$  and let  $\rho(f, g) := \max_{x \in [0, 1]} |f(x) - g(x)|$ . Then  $\rho$  is a metric.
6. Any normed (real or complex) vector space  $(X, \|\cdot\|)$  is a metric space, with respect to the “natural metric”  $\rho(x, y) := \|x - y\|$ ,  $x, y \in X$ , induced by the norm.
7. Let  $(X, \rho)$  be an arbitrary metric space. Any non-empty sub-set  $A \subset X$  of  $X$  equipped with the same metric  $\rho$  is also a metric space.

• *Open and closed sets. Interior and closure of a set.*

Let  $(X, \rho)$  be a metric space. A set  $U \subset X$  is said to be *open* if  $\forall x \in U, B(x, \epsilon) \subset U$ , for some  $\epsilon > 0$ . Here  $B(x, \epsilon)$  denotes the *open ball* with radius  $\epsilon > 0$  centered at  $x$ :

$$B(x, \epsilon) := \{y \in X : \rho(x, y) < \epsilon\}.$$

A set  $F \subset X$  is *closed*, if its complement  $F^c := X \setminus F$  is open, or equivalently, if  $\forall y \notin F, \exists \epsilon > 0$ , such that  $B(y, \epsilon) \cap F = \emptyset$ . By *convention* the empty set  $\emptyset$  and the entire set  $X$  are open (and hence closed). Any open set  $U$  which contains a given point  $x$  of  $X$  is said to be an *open neighborhood* of  $x$ .

Observe that arbitrary unions of open sets are open and, dually, arbitrary intersections of closed sets are closed. Indeed, if  $U_\alpha \subset_{\text{op}} X$ ,  $\alpha \in \mathcal{A}$ , then for any  $x \in U := \cup_\alpha U_\alpha$ , we have that  $x \in U_{\alpha_0}$  for some  $\alpha_0 \in \mathcal{A}$  and hence  $B(x, \epsilon) \subset U_{\alpha_0} \subset U$  which shows that  $U$  is open.

Also, the intersection of a finite number of open sets is open and correspondingly the union of a finite number of closed sets is closed.

The *closure*  $\bar{A}$  and the *interior*  $\langle A \rangle$  of an arbitrary set  $A \subset X$  are then defined as follows:

$$\bar{A} := \cap \{F : A \subset F, F \subset_{\text{cl}} X\} \quad \text{and} \quad \langle A \rangle := \cup \{U : U \subset A, U \subset_{\text{op}} X\}.$$

The *boundary*  $\partial A$  of  $A$  is defined as  $\partial A := \bar{A} \setminus \langle A \rangle$ .

**Examples:**

1. If  $X = \mathbb{R}$  and  $\rho(x, y) = |x - y|$  then clearly  $(a, b)$  is an open set, for any  $a < b$ ,  $a, b \in \mathbb{R}$  and  $[a, b]$  is a closed set.
2. In the context of the previous example,  $\bar{\mathbb{Q}} = \mathbb{R}$  and  $\langle \mathbb{Q} \rangle = \emptyset$ . Also,  $\partial(0, 1] = \{0, 1\}$  and

A set  $A \subset B$  is said to be *dense* in  $B$  if  $\bar{A} \supset B$ . For example, even though the interior of  $\mathbb{Q}$  is empty, the set  $\mathbb{Q}$  is dense in  $\mathbb{R}$  since its closure is the entire real line.

• *Convergence in metric spaces*

A sequence  $x_n \in X$ ,  $n \in \mathbb{N}$  is said to be *convergent* in the metric space  $(X, \rho)$ , if for some  $x \in X$ ,  $\rho(x_n, x) \rightarrow 0$ , as  $n \rightarrow \infty$ . The point  $x$  is said to be the limit of the sequence  $\{x_n\}$ . As in the case of real sequences, one can define the notion of a *limit point* of a sequence.

As for real sequences, a sequence  $x_n \in X$ ,  $n \in \mathbb{N}$  is said to be a Cauchy sequence in the metric space  $(X, \rho)$  (or simply, Cauchy sequence in the metric  $\rho$ ), if

$$\rho(x_n, x_m) \rightarrow 0, \quad \text{as } n, m \rightarrow \infty.$$

**Definition 3.2** A metric space  $(X, \rho)$  is said to be *complete* if any Cauchy sequence in  $(X, \rho)$  has a limit.

**Definition 3.3** A metric space is said to be *separable*, if it has countable dense subset.

Many infinite dimensional linear metric spaces are *not complete*. Most spaces we will encounter however will be complete and separable.

**Examples:**

1.  $\mathbb{R}^k$  equipped with an arbitrary norm  $\|\cdot\|$  is a complete separable metric space with respect to the metric  $\rho(x, y) := \|x - y\|$ ,  $x, y \in \mathbb{R}^k$ , induced by the norm.
2. Consider the set of continuous functions  $f : [0, 1] \rightarrow \mathbb{R}$ , defined on the closed interval  $[0, 1]$ . We will show in Proposition 4.1 below that  $C[0, 1]$ , equipped with the metric  $\rho_\infty(f, g) := \sup_{x \in [0, 1]} |f(x) - g(x)|$  is a complete and separable metric space.

• *Compact sets. Continuous functions.*

We now review the important notion of compactness. Recall first that a collection of sets  $U_\alpha$ ,  $\alpha \in \mathcal{A}$  is said to be a cover of  $F$  if  $F \subset \cup_\alpha U_\alpha$ . A sub-collection  $U_\alpha$ ,  $\alpha \in \mathcal{B}$ , where  $\mathcal{B} \subset \mathcal{A}$  is said to be a sub-cover if it is also a cover of  $F$ . A cover is open if its elements are open sets.

**Definition 3.4** A set  $F \subset \mathbb{R}$  is compact, if and only if every open cover of  $F$  has a finite sub-cover.

Any closed sub-set of a compact set is also compact.

One can prove that in  $\mathbb{R}^k$ , equipped with the usual Euclidean norm, for example, a set  $F \subset \mathbb{R}^k$  is *compact* if and only if it is *closed* and *bounded*. This is not the case for infinitely-dimensional linear normed spaces for example. The Bolzano–Weierstrass theorem however extends to the case of compact sets in an arbitrary complete metric space. Namely:

**Theorem 3.1** Let  $(X, \rho)$  be a complete metric space and  $K \subset X$  be a compact set in  $(X, \rho)$ . Then, any infinite subset  $\{x_n, n \in \mathbb{N}\} \subset K$  has a limit point in  $K$ , that is, there exists  $x \in K$  and an infinite sub-sequence  $n_k \rightarrow \infty$ , such that  $\rho(x_{n_k}, x) \rightarrow 0$ , as  $k \rightarrow \infty$ .

To prove the theorem, it is enough to show that there exists  $x \in K$  such that the set  $B(x, \epsilon) \cap \{x_n, n \in \mathbb{N}\}$  is infinite, for any  $\epsilon > 0$ . Now suppose that one cannot find such an  $x \in K$ . Then, for each  $x$ , there exists  $\epsilon = \epsilon_x > 0$  such that the ball  $U_x := B(x, \epsilon_x)$  contains only finite number  $x_n$ 's. Consider the *open cover*  $U_x$ , indexed by  $x \in K$  and observe that due to compactness, one can take a finite sub-cover. This however leads to a contradiction since each set  $U_x$  contains only finite number of  $x_n$ 's, whereas the number of  $x_n$ 's in  $K$  is infinite.

**Definition 3.5** A function  $f : X_1 \rightarrow X_2$  between two metric spaces  $(X_1, \rho_1)$  and  $(X_2, \rho_2)$  is continuous, if  $f^{-1}(U) \subset_{\text{op}} X_1$ , for any  $U \subset_{\text{op}} X_2$ .

One can show that  $f : X_1 \rightarrow X_2$  is continuous if and only if for any  $x \in X_1$  and  $\epsilon > 0$ , there exists  $\delta > 0$ , such that  $\rho_2(f(x), f(x')) < \epsilon$  for any  $x' \in B(x, \delta)$ . That is, the above definition is equivalent to the *usual* definition of a continuous function.

The next result extends a well-known fact for functions defined on the real line to the case of general metric spaces. Namely, that continuous image of a compact set is compact.

**Proposition 3.1** Let  $f : X_1 \rightarrow X_2$  between two metric spaces  $(X_1, \rho_1)$  and  $(X_2, \rho_2)$  be continuous. If  $K \subset X_1$  is compact, then so is  $f(K) \subset X_2$ .

A function  $f : X_1 \rightarrow X_2$  between two metric spaces  $(X_1, \rho_1)$  and  $(X_2, \rho_2)$  is said to be uniformly continuous on a set  $D \subset X_1$ , if

$$\forall \epsilon > 0, \exists \delta = \delta(\epsilon) > 0, \text{ such that } \rho_2(f(x), f(y)) < \epsilon, \text{ for any } x, y \in D, \rho_1(x, y) < \delta.$$

Observe that  $\delta = \delta(\epsilon)$  does not depend on  $x$  and  $y$ .

**Proposition 3.2** Let  $f : X_1 \rightarrow X_2$  be a continuous function between two metric spaces  $(X_1, \rho_1)$  and  $(X_2, \rho_2)$ . Then,  $f$  is uniformly continuous on any compact  $K \subset X_1$ .

For more details, see e.g. Ch. 7 in Royden (1988).

• *Topology*

We conclude this section by stating the definition of a *topological space*, which generalizes the notion of a *metric space*.

**Definition 3.6** Let  $X$  be a non-empty set. A collection  $\mathcal{T} \subset 2^X$  of sub-sets of  $X$  is said to be a *topology* if:

- i.  $\emptyset, X \in \mathcal{T}$
- ii. if  $U_\alpha \in \mathcal{T}, \alpha \in \mathcal{A}$ , then  $\cup_{\alpha \in \mathcal{A}} U_\alpha \in \mathcal{T}$ .
- iii. for any two  $A, B \in \mathcal{T}$ , we have  $A \cap B \in \mathcal{T}$

The elements of the topology  $\mathcal{T}$  are called open sets. Their complements are said to be closed. The pair  $(X, \mathcal{T})$  is said to be a topological space.

Consider a topological space  $(X, \mathcal{T})$ . Given an  $x \in X$ , any open set  $U \in \mathcal{T}$  containing  $x$  is said to be a *neighborhood* of  $x$ .

An infinite sequence  $x_n \in X, n \in \mathbb{N}$  is said to converge to  $x \in X$  in the topology  $\tau$ , if for any open neighborhood  $U$  of  $x$ , there exists  $n_0 \in \mathbb{N}$ , such that  $x_n \in U$ , for all  $n \geq n_0$ .

**Examples:**

1. (*the discrete topology*) For any non-empty set  $X$ , the collection of all sub-sets of  $X$   $\mathcal{T} = 2^X$  is a topology. Any subset of  $X$  is then both closed and open. This topology is the so-called discrete topology.
2. Any metric space  $(X, \rho)$  is also a topological space where the topology  $\mathcal{T}$  is precisely the collection of all open sets in  $X$ .
3. Let  $X = \overline{\mathbb{R}} := \mathbb{R} \cup \{-\infty\} \cup \{\infty\}$  and let  $\mathcal{I}$  be the collection of all open intervals in  $\overline{\mathbb{R}}$ , e.g.  $(a, b), [-\infty, a)$  and  $(b, \infty]$ , for any  $a, b \in \mathbb{R}$ . Let  $\mathcal{T}$  be the topology generated by  $\mathcal{I}$ , i.e. the minimal topology which contains  $\mathcal{I}$ .

*Intersections and relations between topologies* Fix a non-empty set  $X$ . Given an arbitrary family of topologies  $\mathcal{T}_\alpha$ ,  $\alpha \in \mathcal{A}$  on  $X$ , the *intersection*

$$\mathcal{T} = \bigcap_{\alpha \in \mathcal{A}} \mathcal{T}_\alpha$$

is also a topology.

Let  $\mathcal{T}_1$  and  $\mathcal{T}_2$  be two topologies on  $X$ . We say that  $\mathcal{T}_1$  is *weaker* than  $\mathcal{T}_2$ , if  $\mathcal{T}_1 \subset \mathcal{T}_2$ . That is, if any open set in  $\mathcal{T}_1$  is also open in  $\mathcal{T}_2$ .

One can easily see that if  $\mathcal{T}_1 \subset \mathcal{T}_2$ , then any convergent sequence in the *stronger* topology  $\mathcal{T}_2$  is also convergent in the *weaker*  $\mathcal{T}_1$ .

*Induced topology.* Let  $(X, \mathcal{T})$  be a topological space and let  $A \subset X$ ,  $A \neq \emptyset$ . Then the collection of subsets of  $A$   $\mathcal{T}_A := \{B \cap A : B \in \mathcal{T}\}$  is a topology called the *induced topology* on  $A$  from  $\mathcal{T}$ .

*Continuous functions.* A function  $f : X \rightarrow Y$  between two topological spaces  $(X, \mathcal{T}_X)$  and  $(Y, \mathcal{T}_Y)$  is said to be continuous if  $f^{-1}(U) \in \mathcal{T}_X$ , for any  $U \in \mathcal{T}_Y$ .

*How to construct topologies?*

1. Consider a collection of functions  $\mathcal{F} \subset \{f : f : X \rightarrow Y\}$  between two non-empty sets  $X$  and  $Y$ . Let  $\mathcal{T}_Y$  be a topology on  $Y$ . One is often interested in the *minimal topology* on  $X$  with respect to which the functions in  $\mathcal{F}$  are *continuous*. One can easily see that this topology is generated by

$$\bigcup_{f \in \mathcal{F}} f^{-1}(\mathcal{T}_Y), \quad \text{where } f^{-1}(\mathcal{T}_Y) := \{f^{-1}(U) : U \in \mathcal{T}_Y\}.$$

(Is  $f^{-1}(\mathcal{T}_Y)$  a topology, for any  $f : X \rightarrow Y$ ?)

2. (*base*)
3. (*neighborhood base*)

*Basic classification of topologies*

$T_0$

$T_1$

$T_2$

*Hausdorff*

**Examples:**

1. Let  $(X, \mathcal{T}_X)$  and  $(Y, \mathcal{T}_Y)$  be two topological spaces. If  $\mathcal{T}_X$  is the *discrete topology*, then any function  $f : X \rightarrow Y$  is continuous. On the other hand, if  $\mathcal{T}_X = \{\emptyset, X\}$ , then for a “general topology”  $\mathcal{T}_Y$  only the constant functions are continuous.
2. ...

**Remark.** One can have different metrics on a space  $X$  which define the same topology  $\mathcal{T}$ . Moreover, there are examples where in one metric, the space  $X$  may be incomplete and complete in the other. However, since the topology is the same for both metrics, the convergence of sequences is not affected by the choice of a metric. It is therefore useful to work with the complete metric.

## 4 Functions on $\mathbb{R}^k$ : continuity and differentiability

We review here in more detail facts about functions on the real line  $\mathbb{R}$  or the Euclidean space  $\mathbb{R}^k$ ,  $k \in \mathbb{N}$ . We will benefit from the abstract results given in Section 3.

### • Continuous functions

Recall that a function  $f : D \rightarrow \mathbb{R}$ , defined over  $D \subset \mathbb{R}$  is continuous at  $x_0 \in D$ , if

$$\forall \epsilon > 0, \exists \delta > 0, \text{ such that } |f(x) - f(x_0)| < \epsilon, \forall x \in (x_0 - \epsilon, x_0 + \epsilon) \cap D.$$

A function  $f : D \rightarrow \mathbb{R}$  is said to be continuous if it is continuous at all  $x \in D$ .

The last result implies that any continuous function  $f : [a, b] \rightarrow \mathbb{R}$  achieves its maximum and minimum. The next one, shows that the continuous image of a closed interval is an *entire* closed interval. That is, all *intermediate points* between the maximum and the minimum are visited by  $f$ .

**Theorem 4.1 (for the intermediate value)** *Let  $f : [a, b] \rightarrow \mathbb{R}$  be continuous. Then, the image  $f([a, b])$  is the entire closed interval  $[c, d]$ , where  $c := \min_{x \in [a, b]} f(x)$  and  $d := \max_{x \in [a, b]} f(x)$ .*

PROOF: Indeed, suppose that  $c < y < d$  is such that  $y \notin f([a, b])$ . Since  $F := f([a, b])$  is closed and  $y \notin F$  it follows that  $(y - \epsilon, y + \epsilon) \cap F = \emptyset$ . Let  $U_1 := (c - \epsilon/2, y - \epsilon/2)$  and  $U_2 := (y + \epsilon/2, d + \epsilon/2)$  be two open sets. Clearly,  $U_1 \cap U_2 = \emptyset$  and  $F \subset U_1 \cup U_2$ . Therefore, the open sets  $V_1 := f^{-1}(U_1)$  and  $V_2 := f^{-1}(U_2)$  cover  $[a, b]$  (the  $V_i$ 's are open since  $f$  is continuous).

Now, let  $x_*, x^* \in [a, b]$  be such that  $f(x_*) = c$  and  $f(x^*) = d$ . Then,  $x_* \in V_1$  and  $x^* \in V_2$ . Notice that the sets  $V_1$  and  $V_2$  are disjoint. Without loss of generality, suppose that  $c \leq x_* < x^* \leq d$  and define

$$x_2 := \inf\{x : x > x_*, x \in V_2\}.$$

Observe that  $x_2 \notin V_2$ . Indeed, since  $V_2$  is open  $(x_2 - \delta, x_2 + \delta) \subset V_2$  for some  $\delta > 0$ . Hence  $x_2 - \delta/2 < x_2$  and  $x_2 - \delta/2 \in V_2$ , which violates the definition of  $x_2$ . A similar argument shows that  $x_2 \notin V_1$ . This however is a contradiction since  $x_2 \in [a, b] \subset V_1 \cup V_2$ . We have thus shown that the image  $f([a, b])$  is the entire interval  $[c, d]$ .  $\square$

Recall that a function  $f : D \rightarrow \mathbb{R}$  is *uniformly continuous*, if

$$\forall \epsilon > 0, \exists \delta = \delta(\epsilon) > 0, \text{ such that } |f(x) - f(y)| < \epsilon, \forall |x - y| < \delta, x, y \in D.$$

One can show that any continuous function  $f : [a, b] \rightarrow \mathbb{R}$  on a compact interval is also uniformly continuous. The following is an example of a continuous function  $f : (0, 1] \rightarrow \mathbb{R}$  which is in fact *bounded* but not uniformly continuous:

$$f(x) := \sin(1/x), \quad x \in (0, 1]. \tag{4.1}$$

A function  $f : D \rightarrow \mathbb{R}$  is *absolutely continuous*, if for any  $\epsilon > 0$ , there exists  $\delta = \delta(\epsilon) > 0$ , such that for any  $x_i, y_i \in D$ ,  $x_i < y_i$ ,  $i = 1, \dots, n$ :

$$\sum_{i=1}^n |x_i - y_i| < \delta \quad \text{implies} \quad \sum_{i=1}^n |f(x_i) - f(y_i)| < \epsilon.$$

Notice the *absolute continuity* implies *uniform continuity*. It is however a strictly stronger property. For example, the function  $g : [0, 1] \rightarrow \mathbb{R}$ ,

$$g(x) := \begin{cases} x \sin(1/x) & , x \in (0, 1] \\ 0 & , x = 0 \end{cases} \quad (4.2)$$

is uniformly continuous but not absolutely continuous.

- *The completeness and separability of  $C[0, 1]$  in the uniform metric*

Consider a collection of functions  $f_n : D \rightarrow \mathbb{R}$ ,  $D \subset \mathbb{R}$ . They are said to converge *pointwise* to a limit function  $f : D \rightarrow \mathbb{R}$ , if

$$\lim_{n \rightarrow \infty} f_n(x) = f(x), \quad \text{for all } x \in D.$$

The functions  $f_n$  converge *uniformly* to  $f$ , if the last convergence is valid *uniformly in  $x$* , that is:

$$\forall \epsilon > 0, \exists N = N_\epsilon, \quad \text{such that } \forall x \in D, |f(x) - f_n(x)| \leq \epsilon, n \geq N_\epsilon,$$

where  $N_\epsilon$  does not depend on  $x$ , i.e. we have  $\sup_{x \in D} |f(x) - f_n(x)| \leq \epsilon, \forall n \geq N_\epsilon$ .

#### Examples:

1. Let  $f_n(x) := 1 - x/n$ ,  $x \in \mathbb{R}$  observe that  $f_n$  converges pointwise to  $f(x) \equiv 1$ , but the convergence is not uniform on  $\mathbb{R}$ .
2. Observe that in the previous example the convergence is uniform on any bounded interval.

The next result considers the special case of the space of continuous functions on a compact interval. Recall that such functions are necessarily *uniformly continuous*.

**Proposition 4.1**  $C[0, 1]$  equipped with  $\rho_\infty(f, g) := \|f - g\|_\infty := \sup_{x \in [0, 1]} |f(x) - g(x)|$ , is a complete and separable metric space.

PROOF: We will only show the *completeness*. Let  $f_n \in C[0, 1]$  be a Cauchy sequence. Thus for all  $\epsilon > 0$ , exists  $N_\epsilon > 0$ , such that  $\sup_{x \in [0, 1]} |f_n(x) - f_m(x)| < \epsilon$ , for all  $n, m \geq N_\epsilon$ . Since

$$|f_n(y) - f_m(y)| \leq \sup_{x \in [0, 1]} |f_n(x) - f_m(x)|, \quad \text{for any } y \in [0, 1],$$

the real sequence  $f_n(y)$ ,  $n = 1, 2, \dots$  is Cauchy. Hence, the sequence of functions converges *pointwise* and  $f_n(x) \rightarrow f(x)$ ,  $n \rightarrow \infty$ , for all  $x \in [0, 1]$ .

It remains to show that this convergence is uniform and that the function  $f$  belongs to  $C[0, 1]$ . We will do so by using several lemmas which are of independent interest.

**Lemma 4.1** Let  $g_k \in C[0, 1]$  be such that  $\sum_{k=1}^{\infty} \|g_k\|_\infty < \infty$ . Then, the series  $G_n(x) := \sum_{k=1}^n g_k(x)$  is convergent, uniformly in  $x \in [0, 1]$ , that is,

$$\|G_n - G\|_\infty = \sup_{x \in [0, 1]} |G_n(x) - G(x)| \longrightarrow 0, \quad \text{as } n \rightarrow \infty,$$

where  $G(x) := \sum_{k=1}^{\infty} g_k(x)$ .

PROOF: The fact  $\sum_{k=1}^{\infty} \|g_k\|_{\infty} < \infty$  implies that the series  $G(x) := \sum_{k=1}^{\infty} g_k(x)$  is convergent for all  $x \in [0, 1]$ . Now, observe that for any  $x \in [0, 1]$ , by the triangle inequality, we have

$$|G(x) - G_n(x)| \leq \sum_{k=n+1}^{\infty} |g_k(x)| \leq \sum_{k=n+1}^{\infty} \|g_k\|_{\infty} \rightarrow 0, \quad \text{as } n \rightarrow \infty.$$

The fact the the right-hand side of the last expression does not depend on  $x$  implies the *uniformity* of the convergence.  $\square$

The next lemma shows that  $C[0, 1]$  is *closed* with respect to uniform convergence.

**Lemma 4.2** *Let  $f_n \in C[0, 1]$ . If  $\|f_n - f\|_{\infty} \rightarrow 0$ ,  $n \rightarrow \infty$  then,  $f \in C[0, 1]$ .*

PROOF: Let  $x_0 \in [0, 1]$  and  $\epsilon > 0$ . Observe that

$$|f(x) - f(x_0)| \leq |f(x) - f_m(x)| + |f_m(x) - f_m(x_0)| + |f_m(x_0) - f(x_0)| \leq 2\|f - f_m\|_{\infty} + |f_m(x) - f_m(x_0)|.$$

Thus, pick sufficiently large  $m$  so that  $\|f_m - f\|_{\infty} < \epsilon/3$ . Fix  $m$ , and apply the continuity of  $f_m$  to conclude the proof.  $\square$

We now return to the proof of the completeness of  $C[0, 1]$ . Given a Cauchy sequence  $f_n \in C[0, 1]$ , we construct the following *special* sub-sequence  $n_k \rightarrow \infty$ . For any  $k \in \mathbb{N}$ , let  $n_k$  be such that

$$\|f_{n_{k+1}} - f_{n_k}\|_{\infty} \leq 1/2^k. \quad (4.3)$$

This can be done by using the definition of a Cauchy sequence. Now, observe that

$$f_{n_k}(x) = f_{n_1}(x) + (f_{n_2}(x) - f_{n_1}(x)) + \cdots + (f_{n_k}(x) - f_{n_{k-1}}(x)) = \sum_{j=1}^k g_j(x),$$

where  $g_j(x) := f_{n_j}(x) - f_{n_{j-1}}(x)$  and  $g_1(x) = f_{n_1}(x)$ .

Notice that by (4.3),  $\sum_{j \geq 1} \|g_j\|_{\infty} < \infty$  and hence Lemma 4.1 implies that  $\|f_{n_k} - f\|_{\infty} \rightarrow 0$ , as  $k \rightarrow \infty$ . Lemma 4.2 implies that  $f \in C[0, 1]$ .

We have thus shown that the subsequence  $f_{n_k}$  converges to  $f \in C[0, 1]$  in the uniform metric  $\rho_{\infty}$ . Now, let  $\epsilon > 0$  and observe that

$$\|f_n - f\|_{\infty} \leq \|f_n - f_{n_k}\|_{\infty} + \|f_{n_k} - f\|_{\infty}.$$

By using the fact that  $\{f_n\}$  is a Cauchy sequence and the fact that  $\|f_{n_k} - f\|_{\infty} \rightarrow 0$ ,  $k \rightarrow \infty$ , we can make both terms in the RHS of the last inequality less than  $\epsilon/2$ , for all sufficiently large  $n$ 's. This concludes the proof of the *completeness* of  $C[0, 1]$ .  $\square$

*The separability of  $C[0, 1]$  follows from the next result.*

**Theorem 4.2 (Weierstrass)** *For any  $f \in C[0, 1]$ , there exists a sequence of polynomials with real coefficients  $p_n(x) := a_{n0}x^n + a_{n1}x^{n-1} + \cdots + a_{nn}$ , such that*

$$\|f - p_n\|_{\infty} = \sup_{x \in [0, 1]} |f(x) - p_n(x)| \rightarrow 0, \quad \text{as } n \rightarrow \infty.$$

• *Differentiable functions*

Recall that  $f : (a, b) \rightarrow \mathbb{R}$  is differentiable at a point  $x \in (a, b)$ , if

$$\left| \frac{f(x+h) - f(x)}{h} - f'(x) \right| \rightarrow 0, \quad \text{as } h \rightarrow 0,$$

for some  $f'(x) \in \mathbb{R}$  called the derivative of  $f$  at  $x$ . Observe that differentiability at  $x$  implies continuity. If  $f : (a, b) \rightarrow \mathbb{R}$  is said to be a differentiable function if it is differentiable for all  $x \in (a, b)$ . The derivative  $f'(x)$  of a differentiable function need not be continuous. The function  $f$  is said to be smooth or continuously differentiable if its derivative  $f' : (a, b) \rightarrow \mathbb{R}$  is a continuous function.

If  $f(x)$  has a local maxima (minima) at  $x$  and  $f'(x)$  exists then necessarily  $f'(x) = 0$ .

**Theorem 4.3 (the mean value theorem)** *Let  $f : [a, b] \rightarrow \mathbb{R}$  be continuous and let  $f : (a, b) \rightarrow \mathbb{R}$  be differentiable. Then*

$$f(b) - f(a) = (b - a)f'(\xi), \quad \text{for some } \xi \in [a, b].$$

The proof and more details can be found in Ch. 5 of Rudin (1976). The above result implies that if:

- (i)  $f'(x) = 0$  for all  $x \in (a, b)$ , then  $f(x) = 0, \forall x \in [a, b]$ .
- (ii)  $f'(x) \geq 0$  for all  $x \in (a, b)$ , then  $f \nearrow$ .
- (iii)  $f'(x) \leq 0$  for all  $x \in (a, b)$ , then  $f \searrow$ .

If  $f : (a, b) \rightarrow \mathbb{R}$  is  $n$  times continuously differentiable in  $(a, b)$ . For any  $x, x_0 \in (a, b)$ , we have the Taylor formula:

$$f(x) = f(x_0) + (x - x_0)f'(x_0) + \dots + \frac{1}{(n-1)!}(x - x_0)^{n-1}f^{(n-1)}(x_0) + \frac{1}{n!}(x - x_0)^n f^{(n)}(x_0 + \theta(x - x_0)),$$

for some  $\theta = \theta(x) \in [0, 1]$ .

## 5 Riemann integration

We borrow this presentation from Ch. 6 of Rudin (1976).

Let  $f : [a, b] \rightarrow \mathbb{R}$  be a *bounded function* and let

$$a = x_0 < x_1 < \dots < x_n = b, \quad \text{with } \Delta x_i := x_i - x_{i-1},$$

be a *partition* of the interval with *diameter*  $d(\{x_i\}) := \max_i \Delta x_i$ . The partition  $\{y_j\}_{j=1}^m$  is said to be a *refinement* of  $\{x_i\}$  if

$$\{y_j\}_{j=1}^m \supset \{x_i\}_{i=1}^n,$$

in which case  $d(\{y_j\}) \leq d(\{x_i\})$ .

Introduce the lower and upper Darboux sums, corresponding to  $f$  and a partition  $P = \{x_i\}$ :

$$L(P, f) := \sum_{i=1}^n m_i \Delta x_i \quad \text{and} \quad U(P, f) := \sum_{i=1}^n M_i \Delta x_i,$$

where  $m_i := \inf_{x_{i-1} \leq x \leq x_i} f(x)$  and  $M_i := \sup_{x_{i-1} \leq x \leq x_i} f(x)$ .

**Definition 5.1** The bounded function  $f : [a, b] \rightarrow \mathbb{R}$  is said to be Riemann integrable with integral  $I = \int_a^b f(x)dx$ , if

$$\sup_P L(P, f) = I = \inf_Q U(Q, f),$$

where  $P$  and  $Q$  vary over the set of all partitions of the interval  $[a, b]$ .

Equivalently,  $f$  is Riemann integrable if the Riemann sums

$$S_n(\{\xi_i\}, f) := \sum_{i=1}^n f(\xi_i)\Delta_i, \quad \text{with arbitrary } \xi_i \in [x_{i-1}, x_i],$$

converge to a number  $I := \int_a^b f(x)dx$ , as the diameter of the partition tends to zero:  $d(\{x_i\}) \rightarrow 0$ .

Observe that for any partition  $P = \{x_i\}_{i=1}^n$

$$L(P, f) \leq S_n(\{\xi_i\}, f) \leq U(P, f),$$

and for any refinement  $Q$  of  $P$  i.e.  $P \subset Q$ , we always have

$$L(P, f) \leq L(Q, f) \leq U(Q, f) \leq U(P, f).$$

This is so regardless of whether the bounded function  $f$  is Riemann integrable or not.

Thus intuitively, we have that  $f$  is Riemann integrable if lower and upper Darboux sums have a common limit, as the partitions become finer. Hence, the following criterion (see e.g. Th 6.6 in Rudin (1976))

**Theorem 5.1** *The bounded function  $f : [a, b] \rightarrow \mathbb{R}$  is Riemann integrable iff*

$$\forall \epsilon > 0, \quad U(P, f) - L(P, f) < \epsilon,$$

for some partition  $P = \{x_i\}_{i=1}^n$  of  $[a, b]$ .

Consequently all continuous functions on  $[a, b]$  are Riemann integrable. Indeed, for example, the continuity of  $f : [a, b] \rightarrow \mathbb{R}$  implies its uniform continuity and hence, for all  $\epsilon > 0$ , exist  $\delta > 0$ , such that  $|f(x) - f(y)| < \epsilon$ ,  $\forall x, y : |x - y| < \delta$ . Thus, for any partition  $P$  of diameter less than  $\delta$ :

$$U(P, f) - L(P, f) \leq \sum_{i=1}^n \epsilon \Delta x_i = (b - a)\epsilon.$$

Similarly, one can show that all bounded functions with only finitely many points of discontinuity are also integrable.

- *Riemann–Stieltjes integrals*

Let now  $\alpha : [a, b] \rightarrow \mathbb{R}$  be a monotone non-decreasing. The definition of the Riemann integral of a bounded  $f : [a, b] \rightarrow \mathbb{R}$  extends as follows. For any partition,  $a = x_0 \leq x_1 \leq \dots \leq x_n = b$ , define the Darboux and Riemann sums for  $f$  with  $\Delta x_i$  replaced by  $\Delta \alpha_i := \alpha(x_i) - \alpha(x_{i-1})$ . If the lower and upper sums converge as the diameter of the partition tends to zero, we say that  $f$  has a Riemann–Stieltjes integral

$$\int_a^b f(x)d\alpha(x).$$

For brevity, we use the notation  $f \in \mathcal{R}(\alpha)$  and  $f \in \mathcal{R}$  if  $f$  is Riemann–Stieltjes or Riemann integrable, respectively.

One can show that if  $f$  is monotone (increasing or decreasing) and bounded and  $\alpha$  is continuous, then  $\int_a^b f(x)d\alpha(x)$  exists. Also, if  $f$  and  $\alpha$  have no common points of discontinuity and the bounded  $f$  has only finite number of discontinuities, then  $\int_a^b f(x)d\alpha(x)$  exists.

### Examples and Properties:

1. Let  $p_n \geq 0$  be such that  $\sum_{n=1}^{\infty} p_n < \infty$  and define

$$\alpha(x) := \sum_{n=1}^{\infty} p_n I(x \geq 1/n).$$

Then, for any continuous function  $f : [0, 1] \rightarrow \mathbb{R}$ , we have

$$\int_0^1 f d\alpha = \sum_{n=1}^{\infty} f(1/n)p_n.$$

2. If  $f \in \mathcal{R}(\alpha)$ , then  $|f| \in \mathcal{R}(\alpha)$  and

$$\left| \int_a^b f(x)d\alpha(x) \right| \leq \int_a^b |f(x)|d\alpha(x).$$

3. (*linearity*) For any  $f_1, f_2 \in \mathcal{R}(\alpha)$  and  $a_1, a_2 \in \mathbb{R}$ , we have  $a_1 f_1 + a_2 f_2 \in \mathcal{R}(\alpha)$  and

$$a_1 \int_a^b f_1(x)d\alpha(x) + a_2 \int_a^b f_2(x)d\alpha(x) = \int_a^b (a_1 f_1(x) + a_2 f_2(x))d\alpha(x).$$

4. (*positivity*) If  $f \leq g$  and  $f, g \in \mathcal{R}(\alpha)$ , then

$$\int_a^b f(x)d\alpha(x) \leq \int_a^b g(x)d\alpha(x).$$

5. If  $f \in \mathcal{R}(\alpha_1)$  and  $f \in \mathcal{R}(\alpha_2)$ , then  $f \in \mathcal{R}(\alpha_1 + \alpha_2)$ , and

$$\int_a^b f(x)d\alpha_1(x) + \int_a^b f(x)d\alpha_2(x) = \int_a^b f(x)d(\alpha_1(x) + \alpha_2(x)).$$

For a differentiable  $\alpha : [a, b] \rightarrow \mathbb{R}$ , and any  $f \in \mathcal{R}(\alpha)$ ,  $f \in \mathcal{R}(\alpha)$  iff  $f(x)\alpha'(x) \in \mathcal{R}$ , and in addition

$$\int_a^b f(x)d\alpha(x) = \int_a^b f(x)\alpha'(x)dx, \tag{5.1}$$

(see e.g. Th. 6.17 in Rudin (1976).)

- *Change of variables*

Let  $\varphi : [A, B] \rightarrow [a, b]$  be a strictly increasing and continuous function which is onto (and hence a bijection by monotonicity). Let also  $f : [a, b] \rightarrow \mathbb{R}$  be such that  $f \in \mathcal{R}(\alpha)$  for some non-decreasing  $\alpha$ . Consider the *change of variables*

$$x = \varphi(y) \quad \text{and} \quad \beta(y) := \alpha(\varphi(y)).$$

Then, for  $g(y) := f(\varphi(y))$  we have  $g \in \mathcal{R}(\beta)$  and

$$\int_A^B g(y) d\beta(y) = \int_a^b f(x) d\alpha(x).$$

In particular, with the choice  $\alpha(x) = x$  and  $\beta(y) = \varphi(y)$  differentiable with  $\varphi'(y) \in \mathcal{R}$ , we have in view of (5.1) that

$$\int_{\varphi(a)}^{\varphi(b)} f(\varphi(y)) \varphi'(y) dy = \int_a^b f(x) dx.$$

- *Integration and differentiation*

Let  $f \in \mathcal{R}$ , then  $f$  is also integrable on  $[a, x]$  and one can consider

$$F(x) := \int_a^x f(u) du.$$

Then,  $F : [a, b] \rightarrow \mathbb{R}$  is continuous and moreover, if  $f$  continuous at  $x_0$ , then  $F$  is differentiable at  $x_0$  with  $F'(x_0) = f(x_0)$ .

The *fundamental Theorem of calculus* states that if  $f \in \mathcal{R}$  and for some differentiable function  $F : [a, b] \rightarrow \mathbb{R}$ ,  $F'(x) = f(x)$ ,  $\forall x \in [a, b]$ , then

$$\int_a^b f(x) dx = F(b) - F(a).$$

- *The integration by parts formula*

Let  $F$  and  $G$  be differentiable on  $[a, b]$ , and  $f(x) := F'(x)$  and  $g(x) = G'(x)$  be Riemann integrable on  $[a, b]$ . Then,

$$\int_a^b f(x) G(x) dx = F(b)G(b) - F(a)G(a) - \int_a^b g(x) F(x) dx.$$

- *Improper integrals*

## 6 Miscellanea

*“The perfect is the enemy of the good” (Voltaire (1694 – 1778)).*

We collect here various results, identities and tricks often encountered in statistics. The list is by no means exhaustive.

- *Little “o”, Big “O” notation and asymptotic equivalence*

Let  $f, g : \mathbb{R} \rightarrow \mathbb{R}$  be two real functions. We write

$$f(x) = \mathcal{O}(g(x)), \quad \text{as } x \rightarrow x_0, \quad \text{if } f(x)/g(x) \rightarrow \text{const}, \quad \text{as } x \rightarrow x_0.$$

Also, we write

$$f(x) = o(g(x)), \quad \text{as } x \rightarrow x_0, \quad \text{if } f(x)/g(x) \rightarrow 0, \quad \text{as } x \rightarrow x_0.$$

Here  $x_0 \in [-\infty, \infty]$  can take the values  $\pm\infty$ .

If  $f(x) = \mathcal{O}(g(x))$ ,  $x \rightarrow x_0$ , we say that  $f$  is of the *same order* as  $g$  near  $x_0$  and if  $f(x) = o(g(x))$ ,  $x \rightarrow x_0$ , we say that  $f$  is of *smaller order* than  $g$  near  $x_0$ .

We write

$$f(x) \sim g(x), \quad \text{as } x \rightarrow x_0, \quad \text{if } f(x)/g(x) \rightarrow 1, \quad \text{as } x \rightarrow x_0,$$

and say that  $f$  and  $g$  are *asymptotically equivalent* at  $x_0$ .

**Some examples:**

- $\ln(x) = o(x^{-\delta})$ ,  $x \downarrow 0$ , for all  $\delta > 0$ .
- $\sin(x) = \mathcal{O}(x)$ ,  $x \rightarrow 0$ , and in fact  $\sin(x) \sim x$ ,  $x \rightarrow 0$ .
- $f(x) := (x+1)^\gamma - x^\gamma$ ,  $x > 0$  for some  $\gamma \in \mathbb{R}$ . Then,

$$f(x) = o(x^\gamma), \quad x \rightarrow \infty, \quad \text{but } f(x) = \mathcal{O}(x^{\gamma-1}), \quad x \rightarrow \infty.$$

- Let  $f(x) \sim cx^{-\gamma}$ ,  $x \rightarrow \infty$  for some  $\gamma > 1$  and assume that  $f$  is Riemann integrable on any finite interval. Then, the improper integral  $\int_x^\infty f(u)du$  exists, and

$$\int_x^\infty f(u)du \sim Cx^{-\gamma+1}, \quad \text{as } x \rightarrow \infty.$$

- *Limit identities*
- *Integral identities*

## 7 Exercises

1. Let  $A = \{1, \dots, n\}$ . Show that the set of all subsets of  $A$  has  $2^n$  elements, where the empty set  $\emptyset$  is also counted.

Hint: Consider all possible ordered sequences of  $n$  zeros or ones. Use this construction to represent subsets of  $A$ .

2. Let  $A_\tau \subset X$ , for all  $\tau \in \mathcal{T}$ , where  $\mathcal{T}$  is some non-empty index set. Show the *de Morgan laws*:

$$X \setminus \bigcup_{\tau \in \mathcal{T}} A_\tau = \bigcap_{\tau \in \mathcal{T}} (X \setminus A_\tau) \quad \text{and} \quad X \setminus \bigcap_{\tau \in \mathcal{T}} A_\tau = \bigcup_{\tau \in \mathcal{T}} (X \setminus A_\tau).$$

3. Construct an example of a map  $f : X \rightarrow Y$ , such that  $f(A \cap B) \neq f(A) \cap f(B)$  for some  $A, B \subset X$ .

4. Let  $A, B \subset X \neq \emptyset$ . Show that:

- a.  $I_{A \cap B} = I_A I_B$
- b.  $I_{A \cup B} = \max\{I_A, I_B\}$
- c.  $I_{A \setminus B} = \max\{I_A - I_B, 0\}$
- d.  $I_{X \setminus A} = 1 - I_A$
- e.  $I_{A \Delta B} = |I_A - I_B|$

Where  $I_A$  denotes the indicator function of the set  $A$ .

5. Give a *rigorous* proof of Proposition 1.1 by using the definitions of operations with sets and functions.

Hint: To prove  $A \subset B$  show, take any  $a \in A$  and show that  $a \in B$ . To prove  $A = B$ , show that  $A \subset B$  and  $B \subset A$ .

6. Let  $A_n \subset A_{n+1}$ ,  $n = 1, 2, \dots$  be a monotone increasing sequence of sets. Show that  $\lim_{n \rightarrow \infty} A_n$  exists and that

$$\lim_{n \rightarrow \infty} A_n = \bigcup_{n \in \mathbb{N}} A_n.$$

Suppose that  $A_n \supset A_{n+1}$ ,  $n = 1, 2, \dots$ , is *monotone decreasing* sequence of sets. Does it have a limit? If yes, then identify it and express it in terms of the  $A_n$ 's.

7. Let  $A_n$ ,  $n \in \mathbb{N}$  be an arbitrary sequence of sets. Express the indicator functions of the sets  $A^* := \limsup(A_n)$  and  $A_* := \liminf(A_n)$  in terms of the indicator functions  $f_n(x) := I_{A_n}(x)$  of the  $A_n$ 's.

8. Consider the sequences of sets:

- a.  $A_n := (-1/n, 1 - 1/n^2]$ ,  $n = 1, 2, \dots$
- b.  $A_n := \{\cos(\pi n)\} \cup (-1, 1)$ ,  $n = 1, 2, \dots$

In each case, find the sets  $A^* := \limsup(A_n)$  and  $A_* := \liminf(A_n)$ .

9. Let  $A_n$ ,  $n \in \mathbb{N}$  be a sequence of *disjoint sets* i.e.  $A_n \cap A_m = \emptyset$ ,  $\forall m \neq n$ . Does the limit  $\lim_{n \rightarrow \infty} A_n$  always exist? If yes, then identify it, if not, then construct a counterexample.

10. Let  $\{x_n\}_{n \in \mathbb{N}}$  be a bounded above sequence of real numbers. Set  $x^* := \limsup x_n$ . Let  $\epsilon > 0$  and consider the set

$$A_\epsilon := \{x^*\} \cup \{x_n : x^* - \epsilon < x_n < x^* + \epsilon\}$$

Identify the sets:

$$A_* := \liminf_{k \rightarrow \infty} A_{1/k} \quad \text{and} \quad A^* := \limsup_{k \rightarrow \infty} A_{1/k}.$$

11. Prove that  $x_n$  is a convergent sequence if and only if (2.1) holds.

12. Prove Theorem 2.1.

Hint: Use the method of Cantor.

13. Prove that any Cauchy sequence of real numbers is convergent.

Hint: Show that it is bounded and that it can have only one limit point. Then apply the Bolzano–Weierstrass theorem.

14. Prove the triangle inequality for the Hamming distance defined in (3.1).

15. Show that the set  $A = \{\sin(n) : n \in \mathbb{N}\}$  is dense in  $[-1, 1]$ .

Hint: Show first that the set  $\{a_n := n \bmod 2\pi : n \in \mathbb{N}\}$  is everywhere dense in  $[0, 2\pi]$ . By using the Bolzano–Weierstrass theorem, show first that  $a_{n_k} \rightarrow a_0 \in [0, 2\pi)$ , for some  $a_0$  and for some sequence  $n_k \rightarrow \infty$ . Taking  $|a_{n_{k_1}} - a_{n_{k_2}}| < \epsilon$ , for an arbitrary fixed  $\epsilon > 0$ , show that  $n_{k_1} - n_{k_2} \bmod 2\pi$  is less than  $\epsilon$ . Conclude that the point 0 is also a limit point of the sequence  $\{a_n\}_{n \in \mathbb{N}}$ .

Now, take any  $\epsilon > 0$  and  $0 < a_{m_\epsilon} < \epsilon$ , for some  $m_\epsilon \in \mathbb{N}$ . Notice that one uses here the fact that  $2\pi$  is irrational to ensure that  $a_{m_\epsilon} > 0$ . Now, for an arbitrary  $a \in [0, 2\pi)$ , one has that  $|a_{km} - a| < \epsilon$ , where  $k$  is such that  $k - 1 < a/\epsilon \leq k$ . Conclude that the point  $a$  is also a limit point of  $\{a_n\}$ . Finally, use the uniform continuity of  $x \mapsto \sin(x)$  on  $[0, 2\pi]$ .

16. Use Definition 3.4 to Proposition 3.1.

17. **a.** Show that the function in (4.1) is continuous on  $(0, 1]$  but not uniformly continuous.

**b.** Show that the function in (4.2) is uniformly continuous but not absolutely continuous.

Hint: Consider the values of the functions at the points  $x_k := 2/(2k + 1)\pi$ ,  $k = 1, 2, \dots$

18. Complete the proof of Lemma 4.2.

19. By using Theorem 4.2, prove that  $C[0, 1]$  equipped with the uniform metric is separable.

20. (*The Helly selection theorem*) Let  $f_n : [0, 1] \rightarrow [0, 1]$ ,  $n \in \mathbb{N}$  be a sequence of monotone, non-decreasing functions.

**a.** Show that there exists a sub-sequence  $n_k \rightarrow \infty$ ,  $k \rightarrow \infty$ , such that for all  $x$

$$f(x) := \lim_{k \rightarrow \infty} f_{n_k}(x) \quad \text{exists.} \quad (7.1)$$

That is the sub-sequence of functions  $f_{n_k}$  converges pointwise for all  $x \in [0, 1]$ .

**b.** If the function in part **a.** is *continuous*, then show that the convergence in (7.1) is necessarily uniform in  $x \in [0, 1]$ .

Hint: Prove first that (7.1) holds for a countable dense subset of  $[0, 1]$ . Then, focus on the “jumps” of the limit  $f$  and show that they are at most countably many. In part **b.**, use monotonicity.

## References

Royden, H. L. (1988). *Real Analysis*. Macmillan, third edition.

Rudin, W. (1976). *Principles of Mathematical Analysis*. McGraw–Hill, New York, third edition edition.